How can cow-individual sensor data, national data and drone images improve our understanding of resilience

C. Kamphuis, W. Ouweltjes, M. Poppe, C.A. Mücher and Y. de Haas

Wageningen University and Research, PO Box 9101, 6700HB Wageningen, the Netherlands
Corresponding author: claudia.kamphuis@wur.nl

Abstract

Resilient dairy cows can be characterized by completing multiple lactations, with good (re)productive performances, facing no or few health problems that they overcome easily, and that are efficient and consistent in milk production. Improving resilience has clear advantages, but phenotypic information on this trait is lacking. We explored whether we can fill this gap with information from sensor technologies as these offer high-frequency, continuous, and longitudinal data of individual cows. We did this in three studies: the first study developed cow-specific resilience proxies using milk yield sensors, and correlating these proxies to national data to study herd factors impacting resilience. The natural logarithm transformed variance (LnVAR**)** in daily milk yield appeared an interesting cow-individual resilience proxy: a low LnVAR was genetically correlated to better udder and hoof health, better longevity and fertility, a higher body condition score, and lower ketosis and milk yield. A low LnVar, thus, represents a cow with a good resilience. Subsequently, herd-year effects for LnVar were estimated and correlated with herd performance parameters derived from the national milk recording system. This revealed large differences in resilience between herd-years: the LnVar in the herd-year with the largest effect was >6 times larger than the LnVar in the herd-year with the smallest effect. The positive correlation with the proportion of cows with a rumen acidosis indication (r = 0.31) suggested that feed management may have an important effect on resilience. The second study used sensor data collected during the first lactation to predict a lifetime resilience score using logistic regression or random forest. Both methods had a similar classification performance (accuracy 45-50%). However, random forest required much less data pre-processing to get to this performance. This makes random forest an attractive method to derive information from sensor data, particularly when input becomes even more complex with new sensor technologies entering the market. One of these new technologies could be camera-mounted drones, which were explored in the third study. These drones were used during several field trials, and artificial intelligence was used to detect, locate and identify cows, and obtain specific cow characteristics (height, volume, weight) from the images. Accuracies of >95% for detecting cows, ~91% for identifying cows, and ~88% for obtaining cow characteristics (lying, standing or grazing) were achieved. This makes camera-mounted drones a promising new technology in monitoring traits that can be used for resilience assessment.

Keywords: Resilience, proxies, prediction, sensor technologies.

Introduction

Resilience is the ability of animals to be minimally affected by environmental disturbances, or to quickly recover from them (Colditz *et al.*, 2016). In that perspective, resilient dairy cows can be characterized by completing multiple lactations, with good (re)productive performances, facing no or few health problems that they overcome easily, and that are efficient and consistent in milk production (Adriaens *et al.*, 2020). Improving resilience has clear advantages, e.g., improved animal health and welfare (Mulder and Rashidi, 2017) and reduced antibiotic usage (König and May, 2019). Despite the importance of resilience is clear in itself, breeding or managing for resilience is hampered due to the complexity of the trait and the lack of phenotypic information.

During the past decade, there are fast developments in sensor technologies that are increasingly being adopted on farms. These technologies primarily aim at monitoring specific traits (e.g., milk production), or at facilitating management-by-exception by detecting specific events that require farmer attention, e.g., heat or clinical mastitis. These sensors generate high-frequency, longitudinal and continuous time-series of data of individual cows, and it is this specific characteristic that makes sensors also interesting for phenotyping complex traits such as resilience. More recently, vision and image technologies are increasingly used to retrieve information on, e.g., cow posture for lameness detection (Van Hertem*et al.*, 2018), or for identification of Holstein cattle (Bhole*et al.*, 2019). Both technologies require specific approaches to retrieve the relevant information. For example, regression analyses can be used to analyse the sensor technologies that produce structured data, but this methodology also requires pre-processing to transform raw sensor data into information that can be used by the method itself. Machine learning approaches are increasingly being used (Lokhorst*et al.*, 2019) in the domain of precision dairy farming, but not yet to study resilience. The vision or image technologies require artificial intelligence (deep learning) approaches to retrieve relevant information from the unstructured data. But the feasibility of vision technologies to study resilience has not been explored yet. The usefulness of sensor data to develop cow-specific proxies for resilience has been explored in the current study. We also explored the differences between regression analysis and machine learning (random forest) in predicting lifetime resilience score (as described in Adriaens *et al.*, 2020) using sensor data collected during the first lactation. Finally, we explored the feasibility of camera-mounted drones, in combination with deep learning, to retrieve information of interest to assess resilience in an outdoor situation.

This paper summarizes the main results and lessons we have learned in our attempt to utilize new and commercially available sensors, in combination with national data, in our understanding of resilience.

Developing cow-specific resilience proxies

Daily milk yield recordings were explored for the development of a cow-individual sensor-based proxy for resilience. We used data provided by Cooperation CRV and CRV BV (Arnhem, the Netherlands)from198,745 first-parity Dutch Holstein Frisian cows milked by automatic milking systems. Three steps were subsequently taken: first, individual lactation curves were fitted, using daily milk yield and 4^{th} order 0.7 quantile regression curves, that reflected production potential in absence of disturbances. Second, the natural transformed variance (**LnVar**) of the deviations between measured milk yield and the fitted curve were computed, with the expectation that resilient cows have a smaller range of deviations from their fitted curve, resulting in a lower value for LnVar. Third, LnVar was genetically correlated with health, longevity, fertility, metabolic and production traits. This demonstrated that low LnVar was indeed genetically correlated to better udder and hoof health, better longevity and fertility, a higher body condition score, and lower ketosis at the same level of milk yield. This made LnVar a good cow-individual sensor-based proxy for resilience, which could be used for breeding purposes (Poppe *et al.*, 2020).However, we expected herd management to have an effect on this cow-individual resilience. Therefore, herd-year estimates of the LnVar were assessed for 9,917 herd-year classes based on records of 227,655 primiparous cows from 2,644 herds from the years 2011-2017, while corrected for genetics and year-season effects. These herd-year estimates were considered a herd resilience indicator, expecting that herd-year classes with a low estimate for LnVar contain cows that have, on average, a consistent daily milk yield and are, thus, resilient. Large differences in herd-year estimates were observed (mean 1.34, range 0.38-2.56). Moreover, herd-year estimates of the same herds between years were positively correlated: if a herd had a high LnVar in a certain year, this same herd tended to have high LnVars in other years too. This indicated an effect of management on this trait, and thus, herd-year estimates were related to herd performance parameters derived from the national milk recording system. Herd-year classes with a high LnVar estimate tended to have a high proportion of cows with a rumen acidosis indication (r = 0.31), high somatic cell score (r = 0.19), low fat content (r = -0.18), long calving interval (r = 0.14), low survival to second lactation (r = -0.13), large herd size (r = 0.12), low lactose content (r = -0.12), and high milk production (r = 0.10). These correlations supported the hypothesis that herds with a high LnVar estimate are not resilient. The high correlation with rumen acidosis indication suggests that feed management may play an important role in resilience (Poppe *et al.*, 2021).

Predicting lifetime resilience using sensor data and machine learning

The fluctuations in daily milk production were useful as a proxy for breeding for resilience (Poppe *et al.*, 2020). However, there are more sensor technologies available on farm. This provides the opportunity to combine data from different sensors that are collected early in life to predict lifetime resilience. Adriaens *et al.* (2020) reported on an approach to compute a lifetime resilience score, and they used sensor information to predict this score. They reported that adding features based on activity sensor data improved prediction accuracies significantly ($P < 0.01$), compared to predictions based on daily milk features alone. Poppe *et al.* (2020) and Adriaens *et al.* (2020) have in common that they pre-process the sensor data to retrieve biologically meaningful features to be included in their regression models. However, machine learning approaches, as random forest, are reported to find (non)linear relationships between variables, while requiring little effort for pre-processing (Touw*et al.*, 2013). We studied whether random forests can predict lifetime resilience scores, and whether these algorithms require less pre-processing efforts than a more traditional method like logistic regression. Data for this study originated from Dairy Campus (Lelystad, the Netherlands), and included data of 370 dairy cows that had daily sensor data available for activity, milk yield, rumination, and weight, for at least 100 days during 1-300 DIM of their first lactation. We followed the same approach as Adriaens *et al.* (2020) to compute lifetime resilience scores, and divided the cows into three evenly distributed groups (High, Medium, Low lifetime resilience). Also following Adriaens *et al.* (2020), we aggregated sensor data to daily measurements. We derived 14 features from these daily measurements for each of the four sensor time series, totalling to 56 sensor features. These 56 features were used to develop a stepwise ordinal logistic regression model (considered as reference). We also developed three random forest models: one that used the same features that were deemed significant in the ordinal regression analysis, and one that used all 56 sensor features. The third random forest used the aggregated daily sensor measurements and lactation averages as predictive features.

Table 1. Accuracy of classifying cows for lifetime resilience score.

Model	Accuracy	Range
Logistic Regression	45.1	8.1
Random Forest		
6 features derived from sensor measurements	45.7	8.4
56 features derived from sensor measurements	51.2	10.9
Daily measurements and lactation averages	50.5	6.3

Table 2. Number of cows used in field studies, and performance (in terms of precision) of deep learning models to detect, identify, and characterize poses, height and weight.

	Carus			Juchowo Farm
	2018	2019	2020	2019
Number of cows	4	6	16	100
Performance				
Detecting cows (%)	95.0	96.2		97.3 shaded 99.9 unshaded
Identifying cows (%)	87.6	91.3		--
Characterization of poses (%)	88.7			
Height and Weight (Mean error)		6cm[1]	31kg[2]	

[1]Mean error in height with LIDAR; [2]Mean error with RBG, after removing two extreme outliers.

All models were validated using 10fold cross validation, and the accuracy was used as performance evaluation metric. The accuracy was 45.1% for the logistic regression. The random forest that used the same six features resulted in an accuracy of 45.7%. Using all 56 pre-processed features in a random forest had an accuracy of 51.2%. Lastly, the random forest that used aggregated daily values and lactation averages had an accuracy of 50.5% (Table 1). We concluded that a random forest can reach similar performance as logistic regression, even with less pre-processing efforts (Ouweltjes *et al.*, 2021).

Camera-mounted drones to obtain cow characteristics for resilience

Unmanned Airborne Systems, better known as drones, are an example of a new technology that might be helpful in monitoring cows that are outdoors and further away from the farm, e.g., in rangeland beef production systems. This study aimed specifically at assessing the feasibility of fixed wing and multirotor camera-mounted drones to identify, locate, and retrieve characteristics of cattle that can be of relevance for resilience. Four field studies were conducted on two research farms (Carus, the Netherlands; Juchowo biological farm, Poland) to collect images and video footage by flying the mounted drone over herds on pasture. Material was annotated using a graphical annotation tool called Label Img which uses label object bounding boxes in images. Annotation was done such that analyses could determine whether (1) cows can be detected in the field, (2) individual cows can be identified, (3) cow characteristics can be classified, (4) height and weight can be derived. Analyses were done through deep learning API Nanonets and AgisoftMetshape for imagery, and Yolo for video footages. Results are promising for detection, identification and characterization (Table 2). Detecting cows in the field reached accuracies >95%, where detecting cows in fields without shade reached higher accuracies (99.9%) than in shaded fields (97.3%). Identifying cows reached an accuracy of 91%. This was reached for a small group of animals, each having a distinct coat pattern. Identification of cows that will have the same uniform colour will be challenging with the techniques explored. Characterization of cows into standing, grazing and lying the analyses revealed no difficulties in separating grazing from lying, but separating grazing and standing is challenging. The mean error in estimating height was 6cm with LiDAR, and in estimation of weight 31kg with RGB 3D after removing two extreme outliers. These results suggest that camera-mounted drones could be a promising new technology in monitoring traits that can be used for resilience assessment.

Conclusions

The natural logarithm transformed variance of daily milk is a good proxy for resilience, as a low value of this proxy is genetically correlated to better udder and hoof health, better longevity and fertility, a higher body condition score, and lower ketosis and milk yield. There were large differences in resilience between herds. Correlating with herd-level data from the national milk recording systems revealed a positive correlation with the proportion of cows having a ketosis indication. This indicates that feed management may have an important effect on resilience. Using on-farm sensor data to predict lifetime resilience resulted in similar performance accuracies between regression and random forest. However, the random forest reached this performance with hardly any data pre-processing, in contrast to regression analysis. Finally, camera-mounted drones are a promising approach to locate cows and retrieve cow characteristics that can be related to resilience.

Acknowledgement

The research leading to these results has received funding from European Union's Horizon 2020 research and innovation programme - GenTORE - under grant agreement N° 727213. Also, this work was part of the Breed4Food programme, financed by the Dutch ministry of Economic Affairs (TKI AF-16022), and the Breed4Food partners Cobb Europe, CRV, Hendrix Genetics, and Topigs Norsvin.

List of references

Adriaens, I., N.C. Friggens, W. Ouweltjes, H. Scott, B. Aernouts, and J. Statham. 2020. Productive life span and resilience rank can be predicted from on-farm first-parity sensor time series but not using a common equation across farms. J. Dairy Sci. 103: 7155-7171.

Bohle, A., O. Falzon, and G. Azzopardi. 2019. A computer vision pipeline that uses thermal and RGB images for the recognition of Holstein cattle. In: Proceedings of the International Conference on Computer Analysis of images and patterns. Vento M., and G. Percanella (eds). CAIP 2019, LNCS 11679, p 108-119.

Colditz, I.G., and B.C. Hine. 2016. Resilience in farm animals: Biology, management, breeding and implications for animal welfare. Anim. Prod. Sci. 56: 1961-1983.

Lokhorst, C., R. de Mol, and C. Kamphuis. 2019. Invited review: Big Data in precision dairy farming. Animal. 13: 1519-1528.

König, S., and K. May. 2019. Invited review: Phenotyping strategies and quantitative-genetic background of resistance, tolerance and resilience associated traits in dairy cattle. Animal 13:897-908.

Mulder, H., and H. Rashidi. 2017. Selection on resilience improves disease resistance and tolerance to infections. J. Anim. Sci. 95: 3346-3358.

Ouweltjes, W., M. Spoelstra, D. Ducro, Y. de Haas, and C. Kamphuis. 2021. A data-driven prediction of lifetime resilience of dairy cows using commercial sensor data collected during first lactation. *Under review.*

Poppe, M., R.F. Veerkamp, M.L. van Pelt, H.A. Mulder. 2020. Exploration of variance, autocorrelation, and skewness of deviations from lactation curves as resilience indicators for breeding. J. Dairy Sci. 103: 1667-1684

Poppe, M., H.A. Mulder, C. Kamphuis, and R.F. Veerkamp. 2021. Between-herd variation in and relations to herd performance. J. Dairy Sci. 104: 616-617

Touw, W. G., J.R. Bayjanov, L. Overmars, L. Backus, J. Boekhorst, M. Wels, and S.A. van Hijum. 2013. Data mining in the Life Sciences with Random Forest: a walk in the park or lost in the jungle? Briefings in bioinformatics. 14: 315-326.

Van Hertem, T., A. Schlageter Tello, S. Viazzi, M. Steensels, C. Bahr, C.E. Bites Romanini, K. Lokhorst, E. Maltz, I. Halachmi, and D. Berckmans. 2018. Implementation of an automatic 3D vision monitor for dairy cow locomotion in a commercial farm. Biosystems Engineering. 173: 166-175